J. ULm

INPUT SET: S32159.raw

> **This Raw Listing contains the General**
> **Information Section and up to the first 5 pages.**

```
 1                              SEQUENCE LISTING
 2
 3    (1)    General Information
 4                                                   ENTERED
 5           (i) APPLICANT: Bandman, Olga
 6                               Lal, Preeti
 7                          Shah, Purvi
 8
 9           (ii) TITLE OF THE INVENTION: NEW INTEGRAL MEMBRANE PROTEIN
10
11           (iii) NUMBER OF SEQUENCES: 3
12
13           (iv) CORRESPONDENCE ADDRESS:
14              (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
15              (B) STREET: 3174 Porter Drive
16              (C) CITY: Palo Alto
17              (D) STATE: CA
18              (E) COUNTRY: USA
19              (F) ZIP: 94304
20
21           (v) COMPUTER READABLE FORM:
22              (A) MEDIUM TYPE: Diskette
23              (B) COMPUTER: IBM Compatible
24              (C) OPERATING SYSTEM: DOS
25              (D) SOFTWARE: FastSEQ for Windows Version 2.0
26
27           (vi) CURRENT APPLICATION DATA:
28              (A) APPLICATION NUMBER: 09/265,710
29              (B) FILING DATE:
30              (C) CLASSIFICATION:
31
32           (vii) PRIOR APPLICATION DATA:
33              (A) APPLICATION NUMBER: 08/892,690
34              (B) FILING DATE:
35
36           (viii) ATTORNEY/AGENT INFORMATION:
37              (A) NAME: Billings, Lucy J.
38              (B) REGISTRATION NUMBER: 36,749
39              (C) REFERENCE/DOCKET NUMBER: PF-0339 US
40
41           (ix) TELECOMMUNICATION INFORMATION:
42              (A) TELEPHONE: 415-855-0555
43              (B) TELEFAX: 415-845-4166
44
45              (2) INFORMATION FOR SEQ ID NO:1:
46
```

INPUT SET: S32159.raw

```
47              (i) SEQUENCE CHARACTERISTICS:
48                  (A) LENGTH: 266 amino acids
49                  (B) TYPE: amino acid
50                  (C) STRANDEDNESS: single
51                  (D) TOPOLOGY: linear
52
53              (vii) IMMEDIATE SOURCE:
54                  (A) LIBRARY: BRAINOT03
55                  (B) CLONE: 662708
56
57              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
58
59      Met Val Lys Ile Ser Phe Gln Pro Ala Val Ala Gly Ile Lys Gly Asp
60       1               5                  10                  15
61      Lys Ala Asp Lys Ala Ser Ala Ser Ala Pro Ala Pro Ala Ser Ala Thr
62               20                  25                  30
63      Glu Ile Leu Leu Thr Pro Ala Arg Glu Glu Gln Pro Pro Gln His Arg
64               35                  40                  45
65      Ser Lys Arg Gly Gly Ser Val Gly Gly Val Cys Tyr Leu Ser Met Gly
66           50                  55                  60
67      Met Val Val Leu Leu Met Gly Leu Val Phe Ala Ser Val Tyr Ile Tyr
68       65                  70                  75                  80
69      Arg Tyr Phe Phe Leu Ala Gln Leu Ala Arg Asp Asn Phe Phe Arg Cys
70                       85                  90                  95
71      Gly Val Leu Tyr Glu Asp Ser Leu Ser Ser Gln Val Arg Thr Gln Met
72               100                 105                 110
73      Glu Leu Glu Glu Asp Val Lys Ile Tyr Leu Asp Glu Asn Tyr Glu Arg
74               115                 120                 125
75      Ile Asn Val Pro Val Pro Gln Phe Gly Gly Gly Asp Pro Ala Asp Ile
76           130                 135                 140
77      Ile His Asp Phe Gln Arg Gly Leu Thr Ala Tyr His Asp Ile Leu Asp
78       145                 150                 155                 160
79      Lys Cys Tyr Val Ile Glu Leu Asn Thr Thr Ile Val Leu Pro Pro Arg
80                   165                 170                 175
81      Asn Phe Trp Glu Leu Leu Met Asn Val Lys Arg Gly Thr Tyr Leu Pro
82               180                 185                 190
83      Gln Thr Tyr Ile Ile Gln Glu Glu Met Val Val Thr Glu His Val Ser
84               195                 200                 205
85      Asp Lys Glu Ala Leu Gly Ser Phe Ile Tyr His Leu Cys Asn Gly Lys
86           210                 215                 220
87      Asp Thr Tyr Arg Leu Arg Arg Arg Ala Thr Arg Arg Arg Ile Asn Lys
88       225                 230                 235                 240
89      Arg Gly Ala Lys Asn Cys Asn Ala Ile Arg His Phe Glu Asn Thr Phe
90               245                 250                 255
91      Val Val Glu Thr Leu Ile Cys Gly Val Val
92               260                 265
93
94              (2) INFORMATION FOR SEQ ID NO:2:
95
96              (i) SEQUENCE CHARACTERISTICS:
97                  (A) LENGTH: 1181 base pairs
98                  (B) TYPE: nucleic acid
99                  (C) STRANDEDNESS: single
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/265,710*

INPUT SET: S32159.raw

```
100              (D) TOPOLOGY: linear
101
102          (vii) IMMEDIATE SOURCE:
103              (B) CLONE: 662709
104
105          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
106
107   CGAGCGGGAT CCAAACTTCC GGTGCCTGCA GAGCTCGGAG CGGCGGAGGC AGAGACCGAG      60
108   GCTGCACCGG CAGAGGCTGC GGGGCGGACG CGCGGGCCGG CGCAGCCATG GTGAAGATTA     120
109   GCTTCCAGCC CGCCGTGGCT GGCATCAAGG GCGACAAGGC TGACAAGGCG TCGGCGTCGG     180
110   CCCCTGCGCC GGCCTCGGCC ACCGAGATCC TGCTGACGCC GGCTAGGGAG GAGCAGCCCC     240
111   CACAACATCG ATCCAAGAGG GGGGGCTCAG TGGGCGGCGT GTGCTACCTG TCGATGGGCA     300
112   TGGTCGTGCT GCTCATGGGC CTCGTGTTCG CCTCTGTCTA CATCTACAGA TACTTCTTTC     360
113   TTGCACAGCT GGCCCGAGAT AACTTCTTCC GCTGTGGTGT GCTGTATGAG GACTCCCTGT     420
114   CCTCCCAGGT CCGGACTCAG ATGGAGCTGG AAGAGGATGT GAAAATCTAC CTCGACGAGA     480
115   ACTACGAGCG CATCAACGTG CCTGTGCCCC AGTTTGGCGG CGGTGACCCT GCAGACATCA     540
116   TCCATGACTT CCAGCGGGGT CTGACTGCGT ACCATGATAT CTTGGACAAG TGCTATGTCA     600
117   TCGAACTCAA CACCACCATT GTGCTGCCCC CTCGCAACTT CTGGGAGCTC CTCATGAACG     660
118   TGAAGAGGGG GACCTACCTG CCGCAGACGT ACATCATCCA GGAGGAGATG GTGGTCACGG     720
119   AGCATGTCAG TGACAAGGAG GCCCTGGGGT CCTTCATCTA CCACCTGTGC AACGGGAAAG     780
120   ACACCTACCG GCTCCGGCGC CGGGCAACGC GGAGGCGGAT CAACAAGCGT GGGGCCAAGA     840
121   ACTGCAATGC CATCCGCCAC TTCGAGAACA CCTTCGTGGT GGAGACGCTC ATCTGCGGGG     900
122   TGGTGTGAGG CCCTCCTCCC CCAGAACCCC CTGCCGTGTT CCTCTTTTCT TCTTTCCGGC     960
123   TGCTCTCTGG CCCTCCTCCT TCCCCCTGCT TAGCTTGTAC TTTGGACGCG TTTCTATAGA    1020
124   GGTGACATGT CTCTCCATTC CTCTCCAACC CTGCCCACCT CCCTGTACCA GAGCTGTGAT    1080
125   CTCTCGGTGG GGGGCCCATC TCTGCTGACC TGGGTGTGGC GGAGGGAGAG GCGATGCTGC    1140
126   AAAGTGTTTT CTGTGTCCCA CTGTCTTGAA GCTGGGCCTG C                       1181
127
128          (2) INFORMATION FOR SEQ ID NO:3:
129
130          (i) SEQUENCE CHARACTERISTICS:
131              (A) LENGTH: 263 amino acids
132              (B) TYPE: amino acid
133              (C) STRANDEDNESS: single
134              (D) TOPOLOGY: linear
135
136          (vii) IMMEDIATE SOURCE:
137              (A) LIBRARY: GenBank
138              (B) CLONE: 624778
139
140          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
141
142   Met Val Lys Ile Ala Phe Asn Thr Pro Thr Ala Val Gln Lys Glu Glu
143    1               5                  10                  15
144   Ala Arg Gln Asp Ile Glu Ala Leu Val Ser Arg Thr Val Arg Ala Gln
145               20                  25                  30
146   Ile Leu Thr Gly Lys Glu Leu Arg Val Val Pro Gln Glu Lys Asp Gly
147           35                  40                  45
148   Ser Ser Gly Arg Cys Met Leu Thr Leu Leu Gly Leu Ser Phe Ile Leu
149       50                  55                  60
150   Ala Gly Leu Ile Val Gly Gly Ala Cys Ile Tyr Lys Tyr Phe Met Pro
151   65                  70                  75          -       80
152   Lys Ser Thr Ile Tyr His Gly Glu Met Cys Phe Phe Asp Ser Glu Asp
```

INPUT SET: S32159.raw

```
153                         85                  90                  95
154     Pro Val Asn Ser Ile Pro Gly Gly Glu Pro Tyr Phe Leu Pro Val Thr
155                     100                 105                 110
156     Glu Glu Ala Asp Ile Arg Glu Asp Asp Asn Ile Ala Ile Ile Asp Val
157                 115                 120                 125
158     Pro Val Pro Ser Phe Ser Asp Ser Asp Pro Ala Ala Ile Ile His Asp
159             130                 135                 140
160     Phe Glu Lys Gly Met Thr Ala Tyr Leu Asp Leu Leu Leu Gly Asn Cys
161     145                 150                 155                 160
162     Tyr Leu Met Pro Leu Asn Thr Ser Ile Val Met Thr Pro Lys Asn Leu
163                 165                 170                 175
164     Val Glu Leu Phe Gly Lys Leu Ala Ser Gly Lys Tyr Leu Pro His Thr
165                 180                 185                 190
166     Tyr Val Val Arg Glu Asp Leu Val Ala Val Glu Glu Ile Arg Asp Val
167             195                 200                 205
168     Ser Asn Leu Gly Ile Phe Ile Tyr Gln Leu Cys Asn Asn Arg Lys Ser
169             210                 215                 220
170     Phe Arg Leu Arg Arg Arg Asp Leu Leu Leu Gly Phe Asn Lys Arg Ala
171     225                 230                 235                 240
172     Ile Asp Lys Cys Trp Lys Ile Arg His Phe Pro Asn Glu Phe Ile Val
173                 245                 250                 255
174     Glu Thr Lys Ile Cys Gln Glu
175                 260
176
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/09/265,710*

INPUT SET: S32159.raw

Line   Error   Original Text